EPJ manuscript No.
(will be inserted by the editor)

Spectral rigidity of vehicular streams (Random Matrix Theory approach)

Milan Krbálek[1,2] and Petr Šeba[2,3,4]

[1] Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Prague - Czech Republic,

[2] Doppler Institute for Mathematical Physics and Applied Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Prague - Czech Republic

[3] University of Hradec Králové, Hradec Králové - Czech Republic

[4] Institute of Physics, Academy of Sciences of the Czech Republic, Prague - Czech Republic

Corresponding author e-mail: milan.krbalek@fjfi.cvut.cz

Received: date / Revised version: date

Abstract. Using the methods originally developed for Random Matrix Theory we derive an exact mathematical formula for number variance $\Delta_{\mathtt{N}}(L)$ (introduced in [4]) describing a rigidity of particle ensembles with power-law repulsion. The resulting relation is consequently compared with the relevant statistics of the single-vehicle data measured on the Dutch freeway A9. The detected value of an inverse temperature β, which can be identified as a coefficient of a mental strain of car drivers, is then discussed in detail with the respect to the traffic density ϱ and flow J.

PACS. 05.40.-a Fluctuation phenomena, random processes, noise, and Brownian motion – 89.40.-a transportation – 05.45.-a Nonlinear dynamics and chaos

1 Terminus a quo

The main goal of this paper is to show that the statistical fluctuations of single-vehicle data (in vehicular flows) can be very well predicted by the methods known from Random Matrix Theory where used for describing the statistics of energy levels in quantum chaotic systems. Above that, we intend to demonstrate that the changes of statistical variances in vehicular samples depends not only on macroscopical traffic quantities and three traffic phases (as published in [1] or [2]), but also on psychological characteristics of driver's decision-making process. We will evince

that transitions among the traffic phases (free flow, synchronized flow, and wide moving jam) cause perceptible changes in the mental strain of car drivers.

As reported in Ref. [1], [2], [3], [4] and [5] three traffic phases show substantially different microscopical properties. It was demonstrated in [6], [7], [8], and [10] that the microscopical traffic structure can be estimated with the help of a repulsive potential (applied locally only) describing the mutual interaction between successive cars in the chain of vehicles. Especially, the probability density $\wp(r)$ for the distance r of two subsequent cars (*clearance distribution*) can be derived by means of an one-dimensional gas having an inverse temperature β and interacting by the repulsive potential $V(r) = r^{-1}$ (as discussed in Ref. [7], [6], and [8]). Concretely, denoting $\Theta(x)$ the Heaviside's function

$$\Theta(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leqslant 0, \end{cases}$$

and $\mathcal{K}_\lambda(x)$ the modified Bessel's function of the second kind of order λ (Mac-Donald's function), the clearance distribution of the above-mentioned thermodynamical traffic gas reads as

$$\wp(r) = \mathcal{A}\,\Theta(r)\,\mathsf{e}^{-\frac{\beta}{r}}\mathsf{e}^{-Br}, \tag{1}$$

where

$$B = \beta + \frac{3 - \mathsf{e}^{-\sqrt{\beta}}}{2}, \tag{2}$$

$$\mathcal{A}^{-1} = 2\sqrt{\frac{\beta}{B}}\mathcal{K}_1\big(2\sqrt{B\beta}\big). \tag{3}$$

We remark that $\wp(r)$ fulfils two normalization conditions

$$\int_{\mathbb{R}} \wp(r)\,\mathsf{d}r = 1 \tag{4}$$

and

$$\int_{\mathbb{R}} r\,\wp(r)\,\mathsf{d}r = 1. \tag{5}$$

The latter represents a scaling to the mean clearance equal to one, which is introduced for convenience. The above-mentioned distribution (1) is in a good agreement with the clearance distribution of real-road data ([6], [7], [4], and [8]) whereas the inverse temperature β is related to the traffic density ϱ. We note that the inverse temperature β can be understood as a quantitative description of the mental strain of drivers in a given traffic situation. More specially, the parameter β reflects the psychological pressure level under which the driver is if moving in the traffic stream. Whereas the free flows induces practically no mental strain of drivers, with the increasing traffic density the mental strain escalates rapidly. This will be confirmed in the following text.

2 Statistical variances in traffic data

Another powerful way to inspect the interactions between cars within the highway data is to investigate the traffic flow fluctuations. One possibility is to use the so-called *time-gap variance* Δ_{T} considered in paper [9] and defined as follows. Let $\{t_i : i = 1 \dots Q\}$ be the data set of time intervals between subsequent cars passing a fixed point on the highway. Using it one can calculate the moving average

$$T_k^{(N)} = \frac{1}{N}\sum_{i=k}^{k+N-1} t_i \qquad (k = 1 \dots Q - N + 1)$$

of the time intervals produced by the $N+1$ successive vehicles (i.e. N gaps) as well as the total average

$$\overline{T} = \frac{1}{Q}\sum_{i=1}^{Q} t_i \equiv T_1^{(Q)}.$$

The time-gap variance Δ_{T} is defined by the variance of the sample-averaged time intervals as a function of the sampling size N,

$$\Delta_{\mathrm{T}} = \frac{1}{Q-N+1}\sum_{k=1}^{Q-N+1}\left(T_k^{(N)} - \overline{T}\right)^2,$$

where k runs over all possible samples of $N+1$ successive cars. For time intervals t_i being statistically independent the law of large numbers gives $\Delta_{\mathrm{T}}(N) \propto 1/N$.

A statistical analysis of the data set recorded on the Dutch freeway A9 and published in Ref. [9] leads, however, to different results - see the Figure 1. For the free traffic flow ($\varrho < 15$ veh/km/lane) one observes indeed the expected behavior $\Delta_{\mathrm{T}}(N) \propto 1/N$. More interesting behavior, nevertheless, is detected for higher densities ($\varrho > 35$ veh/km/lane). Here Nishinari, Treiber, and Helbing (in Ref. [9]) have empirically found a power law dependence

$$\Delta_{\mathrm{T}}(N) \propto N^{\epsilon}$$

with an exponent $\epsilon \approx -2/3$, which can be explained as a manifestation of correlations between the queued vehicles in a congested traffic flow.

There is, however, one substantial drawback of this description. The time-gap variance was introduced *ad hoc* and hardly anything is known about its exact mathematical properties in the case of interacting vehicles. It is therefore appropriate to look for an alternative that is mathematically well understood. A natural candidate is the *number variance* $\Delta_{\mathrm{N}}(L)$ that was originally introduced for describing *a spectral rigidity* of energy levels in quantum chaotic systems, i.e. for describing a structure of eigenvalues in the Random Matrix Theory. $\Delta_{\mathrm{N}}(L)$ reproduces also the variances in the particle positions of a certain class of one-dimensional interacting gases (for example Dyson gas in Ref. [12]) and it is defined as follows.

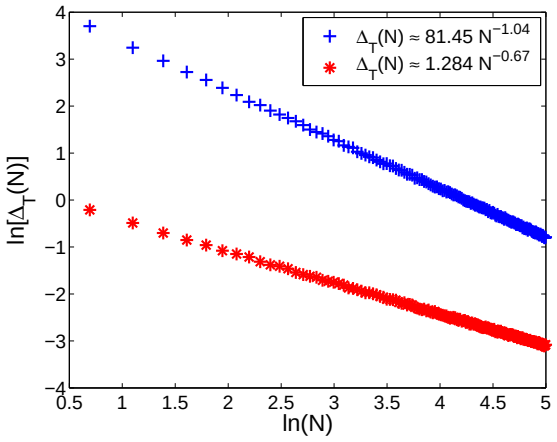

Fig. 1. The time-gap variance $\Delta_{\mathrm{T}}(N)$ as a function of the sampling size N (in log-log scale). Plus signs and stars represent the variance of average time-gaps for free and congested flows, respectively.

Consider a set $\{r_i : i = 1 \ldots Q\}$ of distances (i.e. *clearances* in the traffic terminology) between each pair of subsequent cars moving in the same lane. We suppose that the mean distance taken over the complete set is re-scaled to one, i.e.

$$\sum_{i=1}^{Q} r_i = Q.$$

Dividing the interval $[0, Q]$ into subintervals $[(k-1)L, kL]$ of a length L and denoting by $n_k(L)$ the number of cars

in the kth subinterval, the average value $\overline{n}(L)$ taken over all possible subintervals is

$$\overline{n}(L) = \frac{1}{\lfloor Q/L \rfloor} \sum_{k=1}^{\lfloor Q/L \rfloor} n_k(L) = L,$$

where the integer part $\lfloor Q/L \rfloor$ stands for the number of all subintervals $[(k-1)L, kL]$ included in the interval $[0, Q]$. Number variance $\Delta_{\mathtt{N}}(L)$ is then defined as

$$\Delta_{\mathtt{N}}(L) = \frac{1}{\lfloor Q/L \rfloor} \sum_{k=1}^{\lfloor Q/L \rfloor} \left(n_k(L) - L\right)^2$$

and represents the statistical variance of the number of vehicles moving at the same time inside a fixed part of the road of a length L. The mathematical properties of the number variance are well understood and therefore $\Delta_{\mathtt{N}}(L)$ serves as a better alternative for description of traffic fluctuations (a rigidity of the vehicular chain) than the time-gap variance $\Delta_{\mathtt{T}}(N)$ itself (see also [4]).

3 Exact formula for number variance

Considering the probability density (1) with the only free parameter β (the inverse temperature) we aim to derive an exact formula for the rigidity $\Delta_{\mathtt{N}}(L)$ of thermodynamical traffic gas (see [7]). For these purposes we use the methods presented in [12] and [11] in detail.

Let $\wp_n(r)$ represent the n-th nearest-neighbor probability density (n-th probability density for short), i.e. the probability density for the spacing r of $n+2$ neighboring particles. Owing to the fact that probability density for spacing between two succeeding particles (cars) is $\wp(r) = \wp_0(r)$, the n-th probability density $\wp_n(r)$ can be

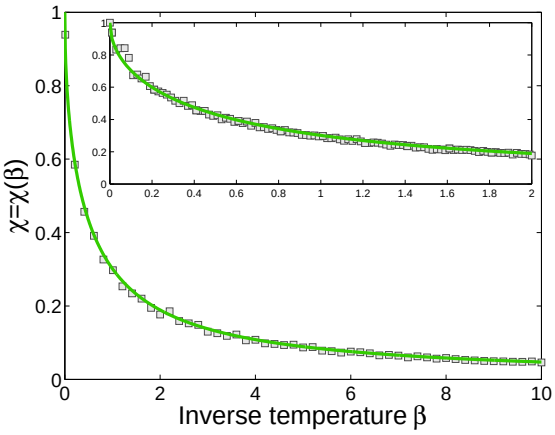

Fig. 2. The slope $\chi(\beta)$ of the number variance $\Delta_{\mathtt{N}}(L) = \chi L + \gamma$ as a function of the inverse temperature β. The squares represent the results of numerical calculations whereas the continuous curve displays the result (8) obtained by the exact mathematical computations. The behavior close to the origin is magnified in the inset.

calculated via recurrent formula

$$\wp_n(r) = \wp_{n-1}(r) \star \wp_0(r),$$

where symbol $\star$ represents a convolution of two independent probabilities, i.e.

$$\wp_n(r) = \int_{\mathbb{R}} \wp_{n-1}(s)\wp_0(r-s)\,\mathtt{d}s.$$

Using the approximation of the function

$$g_n(s) = \mathtt{e}^{-\beta\left(\frac{n^2}{s} + \frac{1}{r-s}\right)}$$

in the saddle point one can obtain

$$\wp_n(r) = \Theta(r)\mathcal{N}_n r^n \mathtt{e}^{-\frac{\beta}{r}(n+1)^2}\,\mathtt{e}^{-Br},$$

where

$$\mathcal{N}_n^{-1} = 2\left(\sqrt{\frac{\beta}{B}}(n+1)\right)^{n+1} \mathcal{K}_{n+1}\left(2(n+1)\sqrt{B\beta}\right)$$

fixes up the proper normalization $\int_{\mathbb{R}} \wp_n(r)\,\mathrm{d}r = 1$. In addition to that the mean spacing equals

$$\int_{\mathbb{R}} r\wp_n(r)\,\mathrm{d}r = n+1.$$

According the book [12] the variance $\Delta_{\mathtt{N}}(L)$ could be evaluated by the formula

$$\Delta_{\mathtt{N}}(L) = L - 2\int_0^L (L-r)\big(1-R(r)\big)\,\mathrm{d}r, \qquad (6)$$

where

$$R(r) = \sum_{n=0}^{\infty} \wp_n(r)$$

is the *two-point cluster function.* The convenient way to calculate the asymptotic behavior of the variance $\Delta_{\mathtt{N}}(L)$ for large L is the application of the Laplace transformation to the two-point cluster function $y(t) = \int_{\mathbb{R}} R(r)\,\mathtt{e}^{-rt}\,\mathrm{d}r$. It leads to a partial result

$$y(t) = \sum_{n=0}^{\infty} \left(\frac{B}{B+t}\right)^{\frac{n+1}{2}} \frac{\mathcal{K}_{n+1}\big(2(n+1)\sqrt{(B+t)\beta}\big)}{\mathcal{K}_{n+1}\big(2(n+1)\sqrt{B\beta}\big)}.$$

The small$-x$ asymptotic behavior of the Mac-Donald's function

$$\mathcal{K}_n(x) \approx \frac{2^{n-1}\Gamma(n)}{x^n}\,\mathtt{e}^{-x} \qquad (x \ll 1)$$

(where $\Gamma(x)$ represents the gamma-function) provides the approximation

$$y(t) = \left(\frac{B+t}{B}\,\frac{\mathtt{e}^{2\sqrt{(B+t)\beta}}}{\mathtt{e}^{2\sqrt{B\beta}}} - 1\right)^{-1}.$$

Applying the Maclaurin's expansion (Taylor's expansion about the point zero) of the function $h(t) = t\cdot y(t)$ to order t^2 we obtain

$$y(t) \approx \frac{1}{t} + \alpha_0 + \alpha_1 t + \mathcal{O}(t^2),$$

where

$$\alpha_0 = -\frac{2B\beta + 3\sqrt{B\beta}}{4\big(1+\sqrt{B\beta}\big)^2},$$

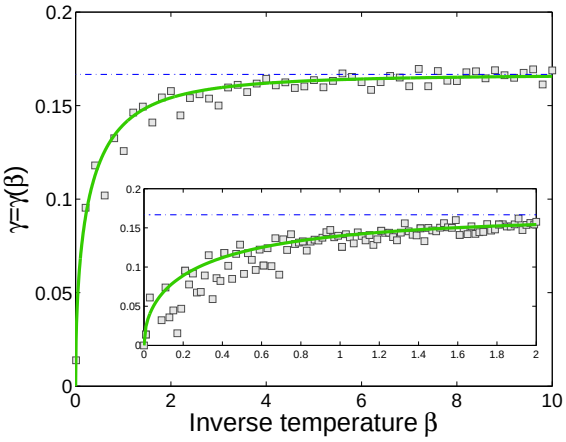

Fig. 3. The shift $\gamma(\beta)$ of the number variance $\Delta_{\mathtt{N}}(L) = \chi L + \gamma$ as a function of the inverse temperature β. The squares represent the results of numerical calculations whereas the continuous curve visualizes the result (9) obtained by the exact mathematical computations. The behavior close to the origin is magnified in the inset. The dash-dotted line displays the asymptotic tendency in $\gamma = \gamma(\beta)$, i.e. $\lim_{\beta\to\infty}\gamma(\beta) = 1/6$.

$$\alpha_1 = \frac{6\sqrt{B\beta} + B\beta\big(21 + 4B\beta + 16\sqrt{B\beta}\big)}{48B\big(1+2\sqrt{B\beta}\big)^3}.$$

Then we get from equation (6)

$$\Delta_{\mathtt{N}}(L) = \chi L + \gamma + \mathcal{O}(L^{-1}), \qquad (7)$$

where

$$\chi = \chi(\beta) = \frac{2+\sqrt{B\beta}}{2B(1+\sqrt{B\beta})} \qquad (8)$$

and

$$\gamma = \gamma(\beta) = \frac{6\sqrt{B\beta} + B\beta\big(21 + 4B\beta + 16\sqrt{B\beta}\big)}{24\big(1+\sqrt{B\beta}\big)^4}. \qquad (9)$$

This crowns the effort to derive an exact form for the number variance. The linear dependence (7) represents a large$-L$ approximation and its correctness is demonstrated in the Fig. 2 and Fig. 3 where compared to the results of numerical computations.

4 Number variance of traffic data

As already discussed the behavior of the number variance is sensitive to the temperature β - or in the terminology of the Random Matrix Theory - to the universality class of the random matrix ensemble. To use the known mathematical results one has not to mix together states with different densities - a procedure known as *data unfolding* in the Random Matrix Theory. For the transportation this means than one cannot mix together traffic states with different traffic densities (as published in [9]) and hence with a different vigilance of the drivers. So we will perform a separate analysis of the data-samples lying within short density intervals to prevent so the undesirable mixing of the different states.

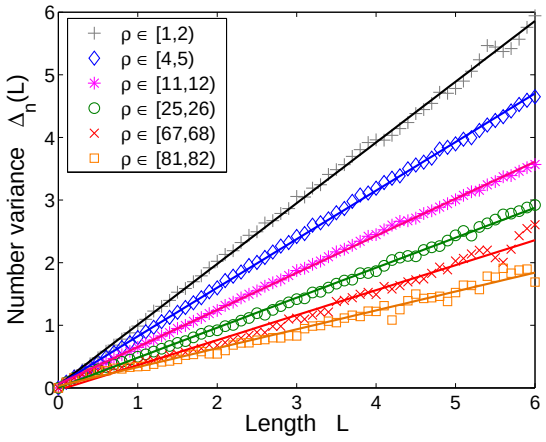

Fig. 4. The number variance $\Delta_{\mathtt{N}}(L)$ as a function of the length L. Plus signs, diamonds, stars, circles, crosses, and squares represent the number variance of real-road data in the selected density regions (see the legend for details). The curves show the linear approximations of the individual data. Their slopes were carefully analyzed and consecutively visualized in the Fig. 5 (top part).

Anyway, we divide the region of the measured densities $\varrho \in [0, 85\,\mathrm{veh/km/lane}]$ into eighty five equidistant subintervals and analyze the data from each one of them separately. The number variance $\Delta_{\mathtt{N}}(L)$ evaluated with the data in a fixed density interval has a characteristic linear tail (see Fig. 4) that is well known from the Random Matrix Theory. Similarly, such a behavior was found in models of one-dimensional thermodynamical gases with the nearest-neighbor repulsion among the particles (see Ref. [11]). We remind that for the case where the interaction is not restricted to the nearest neighbors but includes all particles the number variance has typically a logarithmical tail - see [12]. So the linear tail of $\Delta_{\mathtt{N}}(L)$ supports the view that in the traffic stream the interactions are restricted to the few nearest cars only. The slope of the linear tail of $\Delta_{\mathtt{N}}(L)$ decreases with the traffic density (see the top subplot in the Fig. 5). It is a consequence of the increasing alertness of the drivers and hence of the increasing coupling between the neighboring cars in the dense traffic flows.

The exactly derived properties of the function $\Delta_{\mathtt{N}}(L)$ agree with the behavior of the number variance extracted from the traffic data (see Fig. 4). A comparison between traffic data number variance and the formula (7) allows us to determine the empirical dependence of inverse temperature β on traffic density ϱ. The inverse temperature reflects the microscopic status in which the individual vehicular interactions influence the traffic. Conversely, in the macroscopic approach, traffic is treated as a continuum

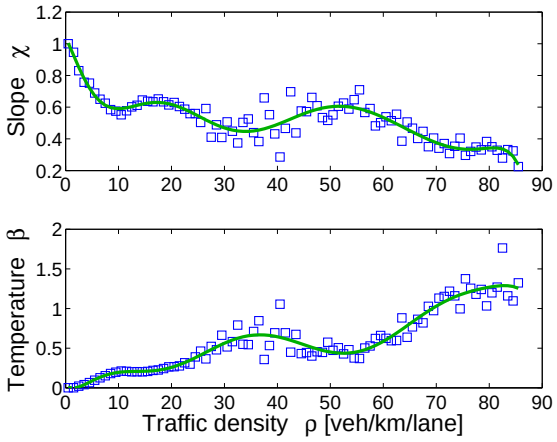

Fig. 5. The slope χ and the inverse temperature β as functions of the traffic density ϱ. The squares on the upper subplot display the slope of the number variance $\Delta_{\mathrm{N}}(L)$ (see Fig. 4), separately analyzed for various traffic densities. The lower subplot visualizes the fitted values of the inverse temperature β, for which the exact form of number variance $\Delta_{\mathrm{N}}(L) = \chi(\beta)\, L + \gamma(\beta)$ corresponds to the number variance obtained from the traffic data. The continuous curves represent polynomial approximations of the relevant data.

and modelled by aggregated, fluid-like quantities, such as density and flow (see [1] and [2]). Its most prominent result is the dependence of the traffic flow on the traffic density - the fundamental diagram.

It is clear that the macroscopic traffic characteristics are determined by its microscopic status. Consequently there should be a relation between the behavior of the fundamental diagram and that of the inverse temperature β. In the Figure 6 we display the behavior of the inverse temperature β simultaneously with the fundamental diagram. The both curves show a virtually linear increase in the region of a free traffic (up to $\varrho \approx 10\,\mathrm{veh/km/lane}$). The inverse temperature β then displays a plateau for the densities up to 18 veh/km/lane while the flow continues to increase. A detailed inspection uncovers, however, that the increase of the traffic flow ceases to be linear and becomes concave at that region. So the flow is reduced with respect to the outcome expected for a linear behavior - a manifestation of the onset of the phenomenon that finally leads to a congested traffic. For larger densities the temperature β increases up to $\varrho \gtrapprox 32\,\mathrm{veh/km/lane}$. The center of this interval is localized at $\varrho \approx 25$ – a critical point of the fundamental diagram at which the flow starts to decrease. This behavior of the inverse temperature is understandable and imposed by the fact that the drivers, moving quite fast in a relatively dense traffic flow, have to synchronize their driving with the preceding car (a strong interaction) and are therefore under a considerable psychological pressure. After the transition from the free to a congested traffic regime (between 40 and 50 veh/km/lane), the synchronization continues to decline because of the decrease in the mean velocity leading to decreasing β. Finally - for densities related to the congested traffic - the inverse temperature increases while the flow remains constant. The comparison between the traffic flow and the inverse temperature is even more illustrative when the changes of the flow are taken into account. Therefore we evaluate the derivative of the flow

$$J' = \frac{\partial J}{\partial \varrho}.$$

The result of the evaluation can be seen from the Figure 5 where one can trace the significant similarities between the shape of $J' = J'(\varrho)$ and inverse temperature β.

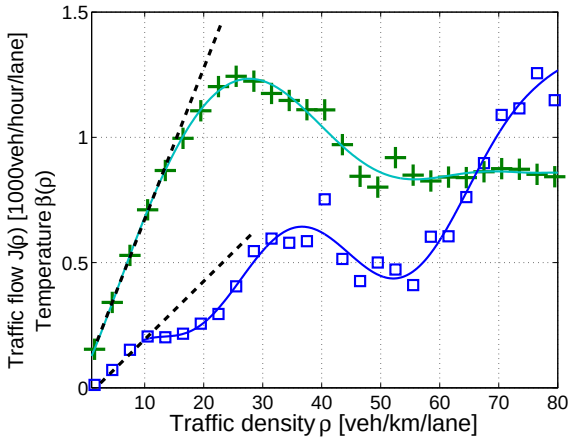

Fig. 6. Traffic flow $J(\varrho)$ and inverse temperature $\beta(\varrho)$ as functions of the traffic density ϱ. Plus signs display a traffic flow in thousands of vehicles per hour and squares correspond to inverse temperature of the traffic gas. The results of a polynomial curve-fitting are visualized by the continuous curves. The dashed lines represent a linear approximations of the relevant data near the origin.

5 Summary and conclusion

With the help of methods of Random Matrix Theory we have derived the exact formula for the rigidity (quantified by the number variance $\Delta_{\mathtt{N}}(L)$) of the one-dimensional thermal particle-gas with the repulsion-potential which depends on the reciprocal value of clearances among the nearest-neighboring particles. The correctness of the result obtained was confirmed by the comparison to the numerical solution.

Analogously to the investigations presented in [4] we have analyzed the number variance $\Delta_{\mathtt{N}}$ of single-vehicle data measured on the Dutch freeway A9. The robust statistical analysis (based on the principle of data unfolding known from Random Matrix Theory) revealed that the rigidity of traffic samples (evaluated by means of the number variance $\Delta_{\mathtt{N}}$) shows a linear dependence

$$\Delta_{\mathtt{N}}(L) \approx \chi L + \gamma \tag{10}$$

(in each of eighty five equidistant density-subregions) whose slope χ depends on traffic density significantly. This knowledge substantially specifies the general results published in Ref. [9].

Above that, the comparison of the exact result (10) with the number variance of traffic data has provided the remarkable possibility how to detect the mental strain level of car drivers moving in the traffic stream. It was demonstrated that the inverse temperature of the traffic sample, indicating the degree of stress of the drivers, shows an increase at both the low and high densities. In the intermediate region, where the free flow regime converts to the congested traffic, it displays more complex behavior.

To conclude, we have shown that microscopical structure of traffic ensembles is rapidly changing with the traffic density and is very well described by the one-parametric class of distributions (1) where the only parameter is the mental strain coefficient β depending on traffic density ϱ. In addition, a recent research (Ref. [13]) shows that a suitable modification (using the knowledge presented in this paper) of Nagel-Schreckenberg cellular model would lead to an interesting progress in the traffic modelling.

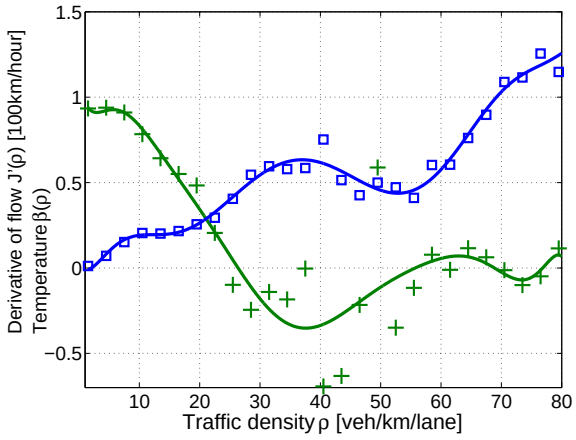

Fig. 7. The inverse temperature $\beta(\varrho)$ and the first derivative $J' = J'(\varrho)$ as functions of the traffic density ϱ. Squares correspond to the inverse temperature of traffic sample while plus signs display the average value of $J' = J'(\varrho)$ (in kilometers per hour). The continuous curves represent relevant polynomial approximations.

Acknowledgements: We would like to thank Dutch Ministry of Transport for providing the single-vehicle induction-loop-detector data. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the projects LC06002 and MSM 6840770039.

References

1. Kerner B S 2004 *The Physics of Traffic* Berlin, New York: Sprinter
2. Helbing D 2001 Rev. Mod. Phys. **73** 1067.
3. Knospe W and Santen L and Schadschneider A and Schreckenberg M 2004 Phys. Rev. E **70** 016115
4. Krbalek M 2008 J. Phys. A: Math. Theor. **41** 205004
5. Kerner B S and Klenov S L and Hiller A and Rehborn H 2006 Phys. Rev. E **73** 046107
6. Krbalek M and Helbing D 2004 Physica A **333** 370
7. Krbalek M 2007 J. Phys. A: Math. Theor. **40** 5813
8. Smith D and Marklof J and Wilson R E 2008 *Preprint* http://rose.bris.ac.uk/dspace/bitstream/1983/1061/1/dasmith.pdf, submitted to Eur. Phys. J. B
9. Helbing D and Treiber M 2003 Phys. Rev. E **68** 067101
10. Helbing D and Treiber M and Kesting A 2006 Physica A **363** 62
11. Bogomolny E B and Gerland U and Schmit C 2001 Eur. Phys. J. B **19** 121
12. Mehta M L 1991 *Random matrices (revised and enlarged)* New York: Academic Press
13. Krbalek M (in preparation)